

Orquestração de Fluxos de Dados com Apache Airflow

Conceitos Básicos - Tasks



No Apache Airflow, "**Tasks**" (Tarefas) referem-se às diversas ações individuais que compõem um fluxo de trabalho completo.

Cada tarefa/task é uma unidade autônoma de execução, representando uma operação específica como a extração de dados de uma fonte, a transformação desses dados ou o carregamento em um destino.

Essas tarefas/tasks são instanciadas a partir de operadores como **PythonOperator**, **BashOperator** ou **PostgresOperator**, que determinam a natureza da operação a ser realizada.

Dentro de um **DAG (Directed Acyclic Graph)**, as tarefas/tasks são interconectadas com base em suas dependências. Isso significa que a execução de uma tarefa/task pode ser condicionada à conclusão bem-sucedida de uma ou mais tarefas anteriores.

Essa estruturação permite que os usuários orquestrem fluxos de trabalho complexos, garantindo que as tarefas/tasks sejam executadas na sequência correta e que as dependências sejam respeitadas, evitando erros e inconsistências.

Além de sua funcionalidade principal, as tarefas no Airflow possuem atributos e métodos que permitem um gerenciamento refinado. Isso inclui definir políticas de reexecução em caso de falhas, configurar notificações, definir prioridades e muito mais.



Muito Obrigado! Continue Trilhando Uma Excelente Jornada de Aprendizagem.